

# SYSTEM AND METHOD FOR CHINESE INPUT USING A JOYSTICK

## CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a Continuation-in-Part of the co-pending U.S. Patent Application No. 10/205,950, titled CHINESE CHARACTER HANDWRITING RECOGNITION SYSTEM filed on 7/25/2002, and claims priority to the U.S. Provisional Patent Application Serial No. 60/503,323, titled TEXT ENTRY METHODS AND SYSTEMS FOR ENTERING CHINESE TEXT USING A JOYSTICK, filed on 16 September 2003, the contents of both of which are incorporated by reference here in their entirety by this reference thereto.

## BACKGROUND OF THE INVENTION

### TECHNICAL FIELD

This invention relates generally to text input technology. More particularly, the invention relates to a system and method that allows a user to use a joystick to input Chinese characters to a data processing device by entering only the first few strokes required to write each character, such that the user can perform Chinese input tasks in a fast, predictive way.

## DESCRIPTION OF THE PRIOR ART

Chinese is written with characters known as *hanzi* ( \* 字). Each character represents a syllable of spoken Chinese and also has a meaning. The characters were originally pictures of people, animals, or other things, but after several thousands years evolution they have become increasingly stylized and no longer resemble the things they represent. Many of the characters are actually compounds of two or more characters. To read Modern Chinese, one should know 2000-3000 characters. To read classical Chinese, knowledge of 5000-6000 characters is required. The largest Chinese dictionaries include 56,000 characters, but most of them are archaic, obscure or in rare variant forms.

A Chinese character (**zi**, 字) can be used as an independent single-character word. It can also be used in combination with other character or characters to constitute a **word** (**ci**, 词) or **phrase** (**cizu**, 词组). In daily life, two-character words are most frequently used. For example, the character “马” means horse and one can make many different words by appending another character to this character:

马(horse)+力(force, power)=马力(horse power);

马(horse)+房(house)=马房(stable, “horse house”);

马(horse)+上(up, on)=马上(immediately, “on horse back”);

马(horse)+夫(man, worker, husband)=马夫(groom, “horse man”);

10 马(horse)+路(road, street)=马路(road, street, “horse path”);

A Chinese character is written using strokes following a specific sequence. There are many different ways to classify Chinese strokes. Traditionally, strokes are classified into eight basic forms as illustrated in FIG. 1. These eight basic strokes can be further grouped into five or six categories depending on different classification criteria.

15 A Chinese character may consist of between 1 and 64 strokes. The strokes for a character are always written in a specific sequence and each stroke is written in a determined direction. In dictionaries, characters are ordered partly by the number of strokes they contain. FIG. 2 shows various sample characters with different number of strokes and a sample four-stroke character with stroke ordinal number and stroke  
20 direction indicated.

In handwriting, it is essential to write a character in the correct sequence. This is not an issue for most of native Chinese speakers because the correct sequence is taught in the first grade and repeated in daily life. There are several basic rules for the sequence,

for example: writing top before bottom, writing left before right, left vertical stroke before top horizontal stroke, bottom horizontal stroke last, center stroke before wings, horizontal strokes before intersecting vertical strokes, left-falling strokes before right-falling strokes, minor strokes (often) last, *etc.*

5 With the broad use of computers, many Chinese input systems and methods have been developed. The systems and methods are primarily classified into two categories: keyboard-coding and handwritten stroke recognition. In a keyboard-coding approach, such as in Three Corners, Goo, 5-stroke (Wubi), or Changjie, the user enters the strokes of a character by pressing on the corresponding key or keys and chooses a  
10 desired character from a set of candidate characters generated and presented on a display as matching alternatives. This approach can also be found in Motorola's CD928C cellular telephone.

In a handwritten stroke recognition approach, such as in Apple-SS, Tegic's T9, or Synaptics' QuickStroke, the user writes a stroke using a special device such as  
15 electronic pen or a stylus and the computer compares the user's stroke with a large number of collections in the database to recognize it. Some of those input systems, such as Zi Corporation's text input solution, have incorporated an intelligent indexing means to intuitively predict and display desired candidates. Some of those systems also include personalization and learning capabilities providing prediction of user-created  
20 terms and frequently used vocabulary.

The purpose of this invention is to provide an alternative Chinese input system and method which uses neither keyboard/telephone keys nor handwriting recognition devices, but uses a joystick or its functional equivalent as a primary input means.

### **SUMMARY OF INVENTION**

25 A Chinese text entry system and method is provided to allow users to enter a character to a device, such as a cellular phone or a PDA, by adding a first few strokes required for the character using a joystick or its functional equivalent. The system and method is

interactive, predictive, and intuitive to use. Compared to other existing solutions, the system and method of the invention can save users considerable time and efforts in Chinese text entry.

By moving the joystick to add one or two strokes which are used to start writing a Chinese character, users can find a desired character from a selection list of characters. The selection list is dynamically created based on the strokes added and the frequency of use of each of the characters starting with the strokes added.

Further, the selection list is context sensitive. When no stroke is added, a default selection list is displayed. The default selection list is typically the first ten most frequently used characters. The sequence of the default selection list changes according to the last character entered. The character that most likely to make a meaningful word by appending to the last character entered is put in the first choice place of the selection list. Similarly, whenever a stroke is added, the corresponding selection list is prioritized by the possibility of making a meaningful word and by the frequency of use.

In the preferred embodiment, the Chinese character input system includes: (1) an input means for associating a joystick motion with one of several basic stroke categories; (2) a collection means for organizing a selection list of characters that commonly start with one or more added strokes, the selection list of characters being displayed in a sequence based on frequency of use/last character entered; and (3) a selection means for selecting and entering the desired character from the selection list.

Typically, the strokes are classified into five basic categories, each of which is represented by a predefined joystick motion.

The collection means contains predefined stroke order information and one or more databases of Chinese characters. It also contains a display means to display a selection list of most frequently used characters when no stroke is entered or immediately after a character is selected and entered. The sequence for the selection list may be based on

any of: (1) use frequency of a character; and (2) contextual relationship between the last character entered and each of the characters in the list.

The selection means includes a means for moving the cursor or its equivalent to select a character from the selection list and a means for entering the selected character. The means for moving the cursor and the means for entering a character can be button pressings or defined joystick motions.

The system may further include means for displaying a numeric or iconic representation of each stroke that is added and a full numeric or iconic representation of strokes for a Chinese character that is entered.

10 According to the preferred embodiment, a method for inputting Chinese characters includes the following steps:

Add a first stroke represented by a predefined joystick motion;

Categorize the added first stroke into one of the predefined number of categories;

15 Find all characters which start with the added first stroke and rank them based on frequency of character use;

Display a list of most frequently used characters of the found characters;

If a desired character is in the list, select the desired character from the list and add it to the composition;

If the desired character is not in the list, add a second stroke;

20 Find most common characters that start with the first stroke and then the second stroke;

Display another list of most frequently characters based on the first and the second stroke; and

Add a third stroke if the desired word is not in the second list.

The method may further comprise the step of:

- 5            Display a default list of characters immediately after a desired character was entered.

The default list is preferably the first ten most frequently used characters. The sequence of the list may be dynamically changed according to the contextual relationship between the last character entered and each of the ten characters.

- 10          The method may further include the steps of:

Display a numeric or iconic representation for a stroke that is added; and

Display full stroke numeric representation for a character that is entered.

### **BRIEF DESCRIPTION OF THE DRAWINGS**

FIG. 1 is a list of basic strokes for Chinese and a number of sample characters;

- 15          FIG. 2 is a list of sample characters with different number of strokes;

FIG. 3 is a schematic, pictorial diagram illustrating a device for inputting Chinese characters using a detached joystick according to the invention;

FIG. 4 is a schematic, pictorial diagram illustrating a device for inputting Chinese characters using an attached joystick according to the invention;

FIG. 5 is a block diagram illustrating the hardware components for a Chinese input system according to the invention;

FIG. 6 is a schematic diagram illustrating the joystick's motion directions in two dimensions;

5 FIG. 7 is a table listing the definitions of the joystick motions and the corresponding strokes and the numeric representations for the strokes;

FIG. 8 is a schematic diagram illustrating an exemplary tree structure of a Chinese vocabulary module;

10 FIG. 9 is a flow diagram illustrating a process for prioritizing the default selection list, e.g. the first ten most frequently used Chinese characters;

FIG. 10 is a flow diagram illustrating a process for prioritizing a dynamically created selection list based on the last entered character and a stroke added;

FIG. 11 is a flow diagram illustrating a method for inputting Chinese characters using a joystick according to the preferred embodiment of the invention;

15 FIG. 12 is a pictorial diagram illustrating an overview of the Chinese input interface before any stroke is added;

FIG. 13 is a pictorial diagram illustrating an overview of the Chinese input interface where a first single horizontal stroke is added;

20 FIG. 14 is a pictorial diagram illustrating an overview of the Chinese input interface where a second horizontal stroke is added;

FIG. 15 is a pictorial diagram illustrating an overview of the Chinese input interface where a third horizontal stroke is added;

FIG. 16 is a pictorial diagram illustrating an overview of the Chinese input interface where a desired character appears to be the first character in the displayed selection list;

5 FIG. 17 is a pictorial diagram illustrating an overview of the Chinese input interface where the first character in the selection list is entered and the default selection list is displayed;

FIG. 18 is a pictorial diagram illustrating an overview of the Chinese input interface where the desired character is not the first character in the selection list and the user moves the joystick to select the desired character;

10 FIG. 19 is a pictorial diagram illustrating an overview of the Chinese input interface where the desired character in FIG. 15 is entered and the default selection list is displayed;

FIG. 20 is a pictorial diagram illustrating an overview of the Chinese input interface where a first stroke for the second desired character is added because after the first  
15 desired character is entered the second desired character was not in the default selection list;

FIG. 21 is a pictorial diagram illustrating an overview of the Chinese input interface where a second stroke for the second desired character is added because the second  
20 desired character was not in the selection list corresponding to the first stroke for the second desired character;

FIG. 22 is a pictorial diagram illustrating an overview of the Chinese input interface where the third stroke is added although the desired character was in the selection list corresponding to the second stroke for the second desired character;



FIG. 23 is a pictorial diagram illustrating an overview of the Chinese input interface where the second desired character is selected by moving the cursor over the character;

FIG. 24 is a pictorial diagram illustrating an overview of the Chinese input interface  
5 where the second desired character is entered and the default selection list is displayed;

FIG. 25 is a pictorial diagram illustrating an overview of the Chinese input interface where a third desired character appears in the default selection list and the third character is selected by moving the cursor over it; and

FIG. 26 is a pictorial diagram illustrating an overview of the Chinese input interface  
10 where a third desired character is entered without adding any stroke and the fourth desired character appears as the first choice in the default selection list.

### **DETAILED DESCRIPTION OF THE INVENTION**

Reference is now made to the drawings, in particular to FIG. 3, which is a schematic, pictorial diagram illustrating a device 101 along with a detached joystick 102 where a  
15 Chinese input system according to the invention embodies. The device 101 has a screen 103 for displaying entered text and a number of buttons, such as the button 104 for various operations. The screen 103 includes a message display area 105 for displaying the entered textual information and a character selection area 106 where a selection list of characters is displayed.

20 The joystick can be incorporated in the main body of the device. FIG. 4 illustrates a portable device 110 in which the Chinese input system embodies. The device 110 includes a joystick 112, a screen 113, and a number of operational buttons, one of which is an ENTER button 114. The screen 113 includes a message display area 115 and a word selection area 116.

The hardware components for the Chinese input system according to the invention are illustrated in FIG. 5, wherein the joystick 122 and the display 123 are communicatively coupled to at least one processor such as the CPU 120 through appropriate interfacing circuitry. Optionally, a speaker 124 may also be coupled to the CPU 120. The user  
5 inputs Chinese characters into the CPU 120 by moving the joystick 122 following a series of pre-defined motions. The CPU 120 manages all output to the display 123, and optionally, to the speaker 124. The system includes a memory 125 which includes a read only random access memory (usually called ROM) for storing application programs and a write/read random access memory (usually called RAM) available for the  
10 application programs' operations. Optionally, the memory 125 also includes an operating system 126 which handles the interface to peripheral hardware, schedules tasks, allocates storage, and presents a default interface to the user when no application program is running. The Chinese input application 127 and the associated Chinese vocabulary modules 128 are responsible for transforming joystick motions into  
15 desired Chinese characters which are displayed on the user's screen.

The Chinese vocabulary modules 128 include at least one database containing a number of Chinese characters and symbols, data of stroke sequence and stroke categories for each single character, data of use frequency of each single character, a  
20 number of multi-character words and phrases, data of use frequency of each multi-character word or phrase, first ten most frequently used characters in a sequence determined by use frequency, the first ten most frequently used characters in a sequence determined by frequency of each of the characters being used as the first character of a sentence or a half-sentence, first ten characters that are most frequently  
25 used characters as the first character of a sentence or a half-sentence, and other data and meta-data for indexing and displaying a character. The database may further include common confusion sets so that the user may enter the alternative entries.

Optionally, the system may include one or more application programs that enable the user to add new phrases to the database for storing multi-character words or phrases.

Optionally, the system may include one or more application programs which count and calculate the user's usage of characters and phrases so that the data of the frequency of use can be dynamically changed. In this way, the more the user uses the system, the more adaptive the system is to the user's need and habit.

- 5    Optionally, the system may also include one or more application programs, such as a word processor 129, a software dictionary 130, and/or a foreign language translator 131. Speech synthesis software may also be provided as an application program which allows the device to function as a communication aid.

10    The joystick 122 is a device consisting of a small handheld stick that pivots about one end and transmits its angle in two dimensions to the CPU 120 via an input/output (I/O) interface. FIG. 6 illustrates various motion directions of the joystick 122. As defined in the table of FIG. 7, when the user moves the joystick 122 from position 0 to position 3, a signal indicating a horizontal stroke category, e.g. — is generated and sent to the CPU 120. Likewise, a motion from position 0 to position 6 represents a vertical stroke category, (e.g. | ); a motion from position 0 to position 7:30 represents a left falling stroke category, (e.g. 丿 ); a motion from position 0 to position 4:30 represents a right falling stroke category, (e.g. ㇏, ㇚); and a motion from position 0 to position 6 then to position 9 represents a hooked stroke category, (e.g. ㇚, ㇚). In addition, a motion from position 0 to position 9 may represent "delete" and a motion from position 0 to position 12 may represent "select."

Alternatively, the joystick 122 can be an isometric joystick in which the input depends on the force exerted rather than the position of the control. In that case, the joystick functions as a TrackPoint.

25    Note that FIG. 7 only shows five basic strokes and their numeric representations according to one criterion for classification. There are other classifications of the Chinese characters. The method and system according to the invention is applicable to any kind of classification.

The display 123 governs the appearance of the Chinese input interface, which includes a message display area and an area for displaying a selection list. Optionally, the display may further include an area for displaying the interface interpretation, either numeric or iconic, of the strokes entered by the user. When a stroke is added, the  
5 numeric or iconic representation of the stroke appears in the area. When a character is entered, the full stroke numeric or iconic representation is displayed. This is useful for assisting users to learn stroke orders and stroke categories.

The Chinese vocabulary modules 128 include a database of Chinese characters which are organized in a tree structure. The root group includes the entirety of the characters,  
10 which are classified into five sub-groups: sub-group 1 includes all characters which start with a “1” stroke; sub-group 2 includes all characters which start with a “2” stroke; sub-group 3 includes all characters which start with the a “3” stroke; sub-group 4 includes all characters which start with a “4” stroke; and sub-group 5 includes all characters which start with a “5” stroke. Each sub-group includes five sub-sub-groups, each of which  
15 corresponds to one of the five categories as its second stroke. Because there are five stroke categories according to the preferred embodiment, each parent node in the tree structure is connected to five child nodes. Each path between a parent node and a child node represents one of the strokes required for a character. FIG. 8 illustrates an exemplary character object tree structure for a Chinese vocabulary module where a  
20 mapping path for a “毛 3115” character is highlighted.

The Chinese vocabulary modules 128 also include a database for storing the first ten most frequently used characters and a large number of two-character words and multi-character words or phrases. This database is used for prioritizing the selection list based on the contextual relationship between the last character entered and any of the  
25 characters in the selection list. The default selection list is the first ten most frequently used characters (“的一不是了我人在有 ▪ ”). Before any stroke is added, or after a punctuation mark such as a period mark (。), a question mark (?), an exclamation mark (!), a semicolon (;) or a comma (,) is entered, the default selection list is in a first default sequence which is made according to the frequency that each of the characters in the

selection is used as the first character of a sentence or half-sentence. For purposes of the discussion herein, a half-sentence is a partial sentence subsequent to a semicolon or a comma. For example, a first default sequence of the default selection list might be "我" - "一不人在有是的了". In operation, the default sequence of the default selection list varies depending on the resources and the method of statistics used.

The default selection list may have a second default sequence which is specifically for the situation where no meaningful contextual relationship between the last character entered and any of the ten most frequently used characters can be established.

Therefore, the default selection list may have three states: (1) the default selection list in a first default sequence which is displayed when no stroke is entered or after a punctuation mark such as a period mark, a question mark, an exclamation mark, a semicolon, or a comma is entered; (2) the default selection list in a second default sequence which is displayed when a meaningful contextual relationship cannot be established between the last character entered and any of the ten most frequently used characters; and (3) the default selection list in a sequence determined according to the contextual relationship between the last character entered and any of the ten most frequently used characters.

Alternatively, the default selection list in its first default sequence may be substituted with a different default selection list which includes ten characters that are most frequently used as the first character of a sentence or a half-sentence according to statistics.

The purpose of the default selection list is to provide the user with the most probable options even before the user adds any stroke. If the desired character is in the selection list, the user just selects it and enters it into the message display area.

The purpose of the default sequence of the default selection list is to provide the user with the most probable character as his first choice as a character to start a sentence or a half-sentence. If the user's desired character happens to be the first choice, because

in the default status the cursor always covers the first choice (which means that the first choice is already selected), the user just needs to make a button pressing or a joystick motion to enter the first choice into the message display area.

5 In a given version of the Chinese vocabulary modules 128, the default selection list and its first two default sequences are not changed. It is directly displayed by the Chinese input application 127 in a proper situation without doing additional calculations. When a desired character is entered, however, the default selection list is displayed in a dynamically created sequence based on the contextual relationship between the last character entered and the characters in the default selection list. FIG. 9 is a flow  
10 diagram illustrating a process for prioritizing the default selection list according to the contextual relationships. The process includes the following steps:

Step 12: Immediately after a desired character is entered, make ten two-character words by appending each of the characters in the default selection list to the last character entered. For example, assuming the last character entered is “好”, the ten  
15 made-up words are: (a) 好的, (b) 好一, (c) 好不, (d) 好是, (e) 好了, (f) 好我, (g) 好人, (h), 好在, (i) 好有, (j) 好 ”. Among these ten words, only (a)(c)(e)(g)(h) are meaningful, and (g)(a)(e)(h) are more frequently used than (c).

Step 14: Look up each of the made-up words against the multi-character words/phrases database.

20 Step 16: Check whether any word is found in the database. In this example, only (a)(c)(e)(g)(h) are meaningful words. They could be found if the database is resourceful enough.

Step 22: If none of the ten made-up words is found in the database, display the default selection list in its second default sequence.

25 Step 18: If any of the ten made-up words is found in the database, prioritize the default selection list according to use frequency. In particular, if all ten made-up words

are found in the database, prioritize the default selection list based on the use frequency of these ten words. If only some of the made-up words are found in the database, the character whose corresponding made-up word has the highest use frequency is placed in the first choice place. The subsequent characters are placed in order according to the use frequency, from high to low, of their corresponding made-up words. Subsequent to these matching characters are the non-matching characters whose made-up words are not found in the database. These non-matching characters are also placed in order based on their use frequency, from high to low. In this example illustrated above, assuming only (a)(c)(e)(g)(h) are found in the database and assuming their use frequencies are in an order of (g)(a)(e)(h)(c), from high to low, according to statistics, the prioritized default selection list would be "人的了在不一是我有 • " instead of the second default sequence "的一不是了我人在有 • ".

Step 20: Display the prioritized default selection list in the selection list display area. If the first choice is the user's desired character, the user just enters it without need to make a selection motion because the first choice is the default selection.

After the first desired character is entered, if the second desired character is not in the default selection list, the user needs to make a joystick motion to add a stroke. When the first stroke is added, the system looks up the database to get the groups of words whose first stroke is the added first stroke, then display the first-stroke-based first ten most frequently used characters in the selection list area. The second desired character may be or may be not in the selection list which is prioritized based on the use frequency of the ten characters. To serve an ordinary user best, the selection list can also be prioritized based both on the contextual relationship and the use frequency as the scheme for prioritizing the default selection list described above. FIG. 10 is a flow diagram illustrating a process for prioritizing the sequence of a dynamically created selection list based on a stroke added. The process includes the following steps:

Step 30: Immediately after a desired character is entered, if the next desired character is not in the default selection list, add a first stroke for the next desired character.

Step 31: Index the group of characters that start with the added stroke and take the first most frequently used characters as candidate selection list.

Step 32: Append each of the ten characters in the candidate selection list to the last character entered to make ten made-up words of combinations. For example, assuming the last character entered is 好, and the candidate selection list based on the first stroke (一) is “一不在有来大到地下要”, the made-up words or combinations are: (1)好一, (2)好不, (3)好在, (4)好有, (5)好来, (6)好大, (7)好到, (8)好地, (9)好下, (10)好要. Among these ten made-up words, only (6)(3)(2)(8) are meaningful, and (6) are more frequently used than (3)(2)(8).

Step 34: Look up each of the made-up words against the multi-character words/phrases database.

Step 36: Check whether there is any word or phrase in the database whose first two characters matches each of the ten made-up words or combinations. In this example, only (6)(3)(2)(8) are meaningful words. They could be found if the database is resourceful enough.

Step 42: If no matching word or phrase is found in the database, the selection list is displayed in a sequence purely based on the use frequency of each of the ten characters.

Step 38: If any multi-character word or phrase whose first two characters matches any of the ten made-up words, prioritize the candidate selection list according to contextual association value and use frequency. The more matching words or phrases a made-up word or combination has in the database, the higher the contextual association value of the made-up word's second character is. In particular, if all ten



made-up words have matching words or phrases in the database, prioritize the candidate selection list based on the use frequency of the matching words or phrases. If only some of the made-up words have matching words or phrases in the database, the character whose corresponding matching word or phrase has the highest use frequency is placed in the first choice place. The subsequent matching characters are placed in order according to the use frequency, from high to low, of their corresponding matching words or phrases. Subsequent to these matching characters are the non-matching characters whose made-up words do not have any matching word or phrase in the database. These non-matching characters are also placed in order based on their use frequency, from high to low. In our example illustrated above, assuming only (6)(3)(2)(8) are found in the database and assuming their use frequencies are in an order of (6)(3)(2)(8), from high to low, according to statistics, the prioritized default selection list would be "大在不地一有来到下要" instead of the default sequence "一不在有来大到地下要".

Step 40: Display the prioritized selection list in the selection list display area. If the first choice is the user's next desired character, the user just enters it without need to make a selection motion because the first choice is the default selection.

The method of prioritizing the dynamically created candidate selection list according to the last character entered and the first stroke added is also applicable in other situations where the added stroke is the second, third, or fourth strokes required for the next desired character. As illustrated in FIG. 8, the more strokes added, the lower the level of the child node containing the desired character. Note that the system according to the invention is not stroke number based. Rather, it is use frequency based. A desired character which requires many strokes may appear in the selection list when no stroke is added or when the first stroke or the second stroke is added. For example, the character "要" consists of nine strokes, but it is available when the first "一" stroke is added; the character "我", one of the first ten most frequently used characters, consists of seven strokes, but it is available even when on stroke is added.

The scheme for prioritizing the selection list based on use frequency does not necessarily satisfy specific users, such as novelists or journalists who always attempt to write creatively using unusual words, but it can best serve the general public because the statistical data for the use frequency is collected from a well-sampled group representative of the general public. For this reason, the Chinese vocabulary modules 128 may have different versions to reflect different needs of different professions or speakers of different dialects. For example, in a medical professional version, the system should be more likely to predict medical words for medical professionals. Similarly, in a Taiwanese version, the system should be more likely to predict the characters frequently used by Taiwanese.

The Chinese input system may further include: the capabilities to match Latin letters, Roman letters, Arabic numerals, punctuation symbols, and emoticons, with predefined sequences of joystick motions; the capability for the user to define his own stroke category or other symbols; optional training session with known characters; optional prompting of the user to clarify between ambiguous stroke interpretations; means to add explicit strokes and/or remedy a stroke misinterpretation; optional indication of level of confidence of stroke interpretations, *e.g.* color-coding each “ink trail” or a smiley-face that frowns when it is uncertain; as well as ability to delete the last stroke(s) in reverse order.

FIG. 11 is a flow diagram illustrating a method for inputting Chinese characters in a predictive manner according to the preferred embodiment of the invention. The method includes the following steps:

Step 50: Add a stroke by a predefined joystick motion as illustrated in the table of FIG. 7. In various deployments, this is done by sending a signal representative of a stroke to the processor.

Step 52: Categorize the added stroke into one of the five categories. This is actually a step of indexing a first level child node in the structure tree of FIG. 8 which matches the first added stroke. For example, when the first stroke “3” is identified, the

whole “group 3” of characters that start with the stroke “3” become candidates for selection.

Step 54: Find characters under the category based on frequency of character use. This can be done by ranking the characters according to a value which is associated with each character and representative of the frequency of use of the character.

Step 56: Display a list of found characters in the word selection area 105 of FIG. 3 (or the word selection area 115 of FIG. 4). The selection list of characters is displayed in a sequence. Preferably, the sequence is based on the use frequency of a found Chinese character and the contextual relation to the prior character entered as illustrated in FIG. 9 and FIG. 10.

Step 58: Check whether the desired character in the selection list.

Step 60: If the desired character is not in the selection list, add next stroke by another joystick motion. For example, as illustrated in FIG. 8, a second stroke “1” is added.

Step 70: If the desired character is in the list, select it by a joystick motion representing a “select” function such as moving from position 0 to position 12 in FIG. 6 or by pressing a button adapted to the joystick. Because the default status of a selection means such as a color cursor is preferably on the first character of the selection list, if the desired word appears in the first choice place, the selection action is unnecessary and the user can directly enter the desired character into the message display area by an “ENTER” action (see Step 72).

Step 72: Put the selected character in the message display area 105 in FIG. 3 (or the message display area 115 in FIG. 4). This is done by a button pressing action such as pressing the “ENTER” button or by a predefined joystick motion.

Step 74: Check whether the message is complete.

Step 76: Add next stroke if the message is not complete.

Step 62 (continued from Step 60 or Step 76): Find most common characters that appear after a previously selected character based on a present stroke order. This also happens before the first stroke, *i.e.* before Step 50. The details of this step have been described in the above paragraphs in connection with FIG. 9 and FIG. 10.

Step 80: Display the selection list, *i.e.* the ten found characters, and the process continues on Step 58.

Optionally, the device 101 in FIG. 3 or 110 in FIG. 4 may have a function to dynamically display the interface interpretation, either numeric or iconic, of the stroke or strokes entered by the user. In that case, the method described above may further comprise the steps of:

Step 90: Display a numeric/iconic representation for a stroke that is added via a joystick motion; and

Step 92: Display a full stroke numeric/iconic representation for a character that is entered.

As an alternative, Step 54 may be replaced by the step of finding characters that commonly start with one or more recognized stroke patterns.

The default selection list may be displayed before step 50 or step 76 as previously described in the paragraphs in connection with FIG. 9.

One of the major advantages of the Chinese input system according to the invention is the great reduction of ambiguities arising in the subtle distinction between certain subtypes of the stroke categories. To reduce ambiguities, we have made further definitions on the subtypes corresponding to the joystick motions. For example, a horizontal motion with a slight upward motion is stroke 1; a horizontal motion with a

slight downward motion is stroke 5; an approximate horizontal motion from position 0 to a position between 2:30 and 3:30 is stroke 1; an approximate vertical motion from position 0 to a position between 11:00 and 1:00 is a select operation, and *etc.*

Often the difference between two joystick motions is too subtle for a processor to differentiate. However, with the user getting more familiar and intuitive to the joystick, the input will be more effective and more efficient.

The operation process illustrated in FIG. 11 is further illustrated in the following paragraphs in conjunction with a series of schematic, pictorial diagrams based on the embodiment of FIG. 4.

FIG. 12 illustrates an overview of the Chinese input interface before any stroke is added. In the character selection area 116, a default selection list in its first default sequence (我 一 不 人 在 有 是 的 了) 301 is displayed. The first default sequence of the default selection list is the sequence of the first ten most frequently used Chinese characters in ordinary usage according to statistics prioritized by the frequency of each of the characters used as the first character of a sentence or a half-sentence. If the user's first desired character happens to be in the default selection list, he just moves the cursor 302, which is a transparent or semi-transparent bitmap, over the desired character to select it, and then makes an ENTER operation. If the first desired character is the first character of the default selection list, the user just enters the character by an ENTER operation without need to make a selection action.

In one deployment, the user moves the cursor for character selection by pressing a "<>" button, and makes an ENTER operation by a joystick motion such as moving from position 0 to position 12.

In another deployment, the user moves the cursor for character selection by moving the joystick, for example, from position 0 to position 12, or from position 10 to position, and makes an ENTER operation by clicking a button.

Yet in another deployment, the user uses the joystick to do both the selection and the ENTER operations.

Yet in another deployment, the selected character automatically entered the message display area right after a short period of time, e.g. 0.5 second, the character was  
5 selected, *i.e.* the cursor moves over on it.

If the desired character is not in the default selection list, the user adds a stroke by moving the joystick. Optionally, a numeric representation for a stroke added is displayed in the stroke number area 303. In the step illustrated in FIG. 12, because no stroke has been added, the area shows a zero. Alternatively, an iconic representation for a stroke  
10 added can be displayed.

Assuming the user adds a first single horizontal stroke, as illustrated in FIG. 13, by a joystick motion 304, *i.e.* from position 0 to position 3. The stroke category is determined to be "1," which is listed in the stroke number area 303. The system looks up the database and displays a first selection list containing the first most frequently used  
15 characters that start with the first added stroke (一不在有来大到地下要) in the selection list display area 116.

Assuming the first desired character is not in the first selection list in FIG. 13, the user adds a second horizontal stroke as illustrated in FIG. 14. After a second horizontal stroke is entered by the joystick motion 305, the system looks up the database and  
20 displays a second selection list containing the first ten most frequently used characters that start with the first added stroke and then the second added stroke (二天无于 ▪ 三 ▪  
▪ ▪ 理). Because the user has added two first category strokes, the stroke number area 303 shows "11."

Assuming further that the desired character is not in the second selection list in FIG. 14,  
25 the user adds a third horizontal stroke a third horizontal joystick motion of the joystick. After the third horizontal stroke is added, as illustrated in FIG. 15, the system looks up

the database and display a third selection list containing the first most frequently used characters that start with the first added stroke and then the second added stroke and then the third added stroke (三春帮丰奏奉秦慧邦 \* ). In the step of FIG. 15, because the user has added three horizontal strokes, the stroke number area 303 shows "111."

- 5 Assuming the user wants to enter the first character in the third selection list, because the first choice is selected by default, *i.e.* the first character in the selection list is always covered by the cursor 301, the user may simply press the ENTER button as illustrated in FIG. 16. Alternatively, the ENTER operation can be done by a defined joystick motion.
- 10 FIG. 17 illustrates the Chinese input interface when the first character (三) in the third selection list is added to the message display area 115. When the desired character is added, the system makes a number of calculations as illustrated in FIG. 9 and displays the default selection list based on the contextual relationship between the last character entered (三) and any of the ten characters in the default selection list (的一不是我人在有 \* ).
- 15 Among the first ten most frequently used characters (的一不是我人在有 \* ), the character "人" is the most possible character that can make a meaningful word with the last character entered (三). If the first character "人" in the default selection list happens to be the user's next desired character, the user can enter it by giving an ENTER command without need to make a selection action because the first character in the selection list is selected by default. Before the user adds another stroke, the stroke
- 20 number area 303 shows the stroke numbers (111) required for the last character entered (三).

- Now assuming the user's first desired character is not the first character in the selection list 301 shown in FIG. 15, and assuming the first desired character is the eighth
- 25 character (慧), the user may, as illustrated in FIG. 18, first move the cursor 302 over the character by a joystick motion such as the motion 308, then enter the character by giving an ENTER command such as a button pressing or a defined joystick motion. The

strength of this system is its predictive abilities. If a user desires a very complex, but somewhat common, character pointed to in the above illustration, the user needs not complete the stroke for that character. As soon as it is displayed in the selection list 301, it can be selected and entered by one or two joystick motions.

- 5 Note that using the joystick, the complex character (慧) with fifteen strokes can be added into the message display area 115 with only four simple movements, *i.e.* three horizontal strokes 111, one for moving the cursor to select the desired character if the desired character is not the first character in the selection list, and one button pressing, or a defined joystick motion, for entering the selected character. Additionally, the entire  
10 stroke order for the character entered is displayed in the stroke number area 303. If the user was used to an alternate stroke order for the character, he can learn the Government Standard stroke order used by this system.

- Once the desired character (慧) is selected and is added to the message display 115 area, as illustrated in FIG. 19, the system displays the default selection list (人的一不是  
15 了我在有 \* ) which is prioritized based on the contextual relationship between the last character entered (慧) and any of the ten most frequently used characters (的一不是了我人在有 \* ). Assuming the next desired character is “孩” which is not in the default list, the user needs to make a joystick motion corresponding to the first stroke (乚) of the next desired character (孩).

- 20 When the first stroke (乚) for the next desired character (孩) is added by a joystick motion 309, as illustrated in FIG. 20, the system looks up the database and displays a selection list containing the first ten most frequently used characters that start with the added stroke (乚). Assuming the last character entered (慧) cannot make up a meaningful word with any of the ten characters, the selection list is prioritized purely  
25 based on the use frequency of each single character. Because the next desired character (孩) is not in the selection list 301 triggered by the joystick motion 309, the



user needs to make another joystick motion corresponding to the second stroke ( 丿 ) for the character (孩).

FIG. 21 illustrates the Chinese input interface when the second stroke ( 丿 5) is added by the joystick motion 310. The desired character (孩) appears in the selection list 301, which is the second character from the right. The user may select the character and enter it into the message display area by one or two simple motions. For example, the user first moves the cursor over the character to select it and then makes an ENTER operation to enter it into the message display area. Assuming the user wants to continue to demonstrate the disambiguation feature of the system by adding the third stroke ( 一 1) of the character (孩), a new selection list triggered by the third stroke is displayed as shown in FIG. 22.

After the joystick motion 311 for the third stroke ( 一 1) is added, the selection list 301 contains two characters (the first 子 and the third 孑 from the left) that are only slightly different from each other. In fact, these two characters have exactly the same stroke order, and choosing from the selection list is the only way to disambiguate the two characters. Note that the third character (孑) is not only less commonly used than the first one (子), but also of a slightly more complex structure.

Note that the second desired character was first visible after the second stroke ( 丿 5) was added (see FIG. 21), and is still a likely choice in the selection list (the second from the left) after the third stroke ( 一 1) is added. If the desired character was removed from the selection list for some reason, it is an indication that the stroke order entered by the user does not match the Government Standard stroke order used in the system. Now as illustrated in FIG. 23, the user selects the second character (孩) by making a joystick motion 312.

Then, as illustrated in FIG. 24, the user enters the selected character (孩) into the message display area. The entered character (孩) is composed of nine strokes. We

selected it at the third stroke motion (ㄣ 551), but could have selected it at the second stroke motion (See FIG. 21).

After the second desired character (孩) is entered, the selection list returns to the default status. Assuming the third desired character is “这”, which appears in the default selection list (the first from the right) as illustrated in FIG. 25, there is no need to make a stroke motion because the ten most frequently used characters are displayed even before the user touches the joystick. The user just selects the character by moving the cursor over the desired character by the joystick motion 313 and enters the character into the message display area 115 by an ENTER operation.

- 10 After the third desired character (这) is entered, the default selection list 301 based on the last character entered (这) is displayed as shown in FIG. 26. The first choice character (人) of the selection list happens to be the fourth desired character. Therefore, the user may enter it by giving an ENTER command.

- 15 The method described above may be carried out by a computer usable medium containing instructions in computer readable form. In other words, the method may be incorporated in a computer program, a logic device, mobile device, or firmware and/or may be downloaded from a network, e.g. a Web site over the Internet. It may be applied in all sorts of text entry.

- 20 Although the invention is described herein with reference to some preferred embodiments, one skilled in the art will readily appreciate that other applications may be substituted for those set forth herein without departing from the spirit and scope of the present invention.

Accordingly, the invention should only be limited by the Claims included below.